

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Wallis, Nicola G.
Shilling, Lisa K.
Mooney, Jeffrey L.
Debouck, Christine
Zhong, YiYi
Jaworski, Deborah D.
Wang, Min
Throup, John P.

(ii) TITLE OF THE INVENTION: Histidine Kinase

(iii) NUMBER OF SEQUENCES: 6

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Dechert, Price & Rhoads
(B) STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
(C) CITY: Philadelphia
(D) STATE: PA
(E) COUNTRY: USA
(F) ZIP: 19103-2793

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:

(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Falk, Stephen T
- (B) REGISTRATION NUMBER: 36,795
- (C) REFERENCE/DOCKET NUMBER: GM10127

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 215-994-2488
- (B) TELEFAX: 215-994-2222
- (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2201 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TAATTTAAAA AGCAACTATT GTATAGAAAA ATACAAAATT TAAAATATAT TACCTTATTA

60

GAAAAAGTCG TAATATGAGG TGTACAAATG ACGCAAATTT TAATAGTAGA AGATGAACAA

120

AACTTAGCAA GATTCTTGA ATTGGAACTC ACACATGAAA ATTACAATGT GGACACAGAG

180

TATGATGGAC AAGACGGTTT AGATAAAGCG CTTAGCCATT ACTATGATT AATCATATTA

240

GATTTAATGT TGCCGTCAAT TAATGGCTTA GAAATTGTC GCAAAATTAG ACAACAAACAA

300

TCTACACCTA TCATTATAAT TACAGCGAAA AGTGATACGT ATGACAAAGT TGCTGGCTT

360

GATTACGGTG CAGACGATTA TATAGTTAAG CCGTTGATA TTGAAGAACT TTTAGCAAGA
420
ATTCGTGCAA TTTTACGTCG TCAGCCACAA AAGGATATTA TCGATGTCAA CGGTATTACA
480
ATTGATAAGA ACGCTTTAA AGTGACGGTA AATGGCGCAG AAATTGAATT AACAAAAACA
540
GAGTATGATT TACTATATCT TCTAGCTGAA AATAAAAACC ATGTTATGCA ACGGAAACAA
600
ATTTTAAATC ATGTATGGGG TTATAATAGT GAAGTAGAAA CAAATGTCGT AGATGTTTAT
660
ATAAGATATT TACGAAACAA GTAAAACCA TACGATCGTG ACAAAATGAT TGAAACAGTT
720
CGTGGCGTTG GGTATGTGAT ACGATGACAA AACGTAATT GCGCAATAAC TGGATTATTG
780
TTACCACGAT GATTACGTTT GTCACGATAT TTTTGTGTTG TTTAATTATT ATTTTTTCT
840
TGAAAGATAC ACTGCATAAT AGTGAGCTTG ATGATGCAGA ACGAAGCTCA AGCGATATTAA
900
ATAATTATTTC TCATTCTAAG CCTGTTAAAG ATATATCTGC ATTAGACTTG AATGCATCTT
960
TAGGTAATTT TCAAGAGATA ATTATTTATG ATGAGCATAA TAATAAATTAA TTTGAGACAT
1020
CGAATGATAA CACAGTGAGA GTTGAACCAG GTTATGAACA CCGTTATTT GACCGCGTAA
1080
TAAAAAAACG CTATAAAGGC ATTGAATATT TAATTATTAA AGAACCAATT ACAACGCAAG
1140
ATTCAAAGG GTATAGCTTG TTAATTCTT CACTAGAAAA TTATGATAAC ATCGTAAAT
1200
CATTGTATAT CATTGCGCTG GCATTTGGAG TGATTGCAAC AATTATAACT GCCACAATCA
1260
GTTATGTATT TTCAACACAA ATTACTAAC CGCTTGTCAAG TTTATCAAAT AAAATGATTG
1320
AGATTCGACG AGATGGTTT CAAAATAAT TGCAATTAAA TACAAATTAT GAAGAAATAG
1380
ATAATTCTGC AAATACGTTT AATGAGATGA TGAGCCAAAT TGAAGAATCA TTTAATCAAC
1440
AAAGACAATT TGTTGAAGAT GCGTCACATG AATTACGAAC ACCATTACAA ATTATTCAAG
1500
GTCATTAAA TTTGATTCAAG CGATGGGAA AAAAGACCC AGCAGTATTA GAAGAATCGT
1560

TAAATATTTC TATTGAAGAA ATGAATCGTA TCATAAAATT AGTCGAAGAA TTACTTGAAT
1620
TGACTAAAGG AGATGTAAAT GACATTCTT CTGAAGCGCA GACCGTGCAT ATTAATGATG
1680
AAATTGCTC GCGAATACAC TCATTAAAAC AATTGCATCC TGATTATCAA TTTGATACGG
1740
ATCTGACATC TAAAAATCTA GAAATTAAAA TGAAACCTCA TCAATTGAA CAATTATTT
1800
TAATCTTAT TGATAATGCA ATCAAATATG ATGTGAAGAA TAAGAAAATT AAAGTTAAGA
1860
CAAGGTTAAA AAATAAGCAA AAAATAATTG AAATTACAGA TCATGGAATT GGTATTCCAG
1920
AGGAAGATCA AGATTCATT TTTGATCGCT TTTATCGAGT GGATAAAATCT CGTTCAAGAA
1980
GTCAAGGCAG TAATGGACTC GGATTATCTA TTGCTAAAAA AATCATTCAA TTAAACGGAG
2040
GATCGATTAA AATTAAAAGT GAAATTAATA AAGGAACAAAC GTTTAAAATC ATATTTAAT
2100
CATGTCTGAG ACGTCAATCA AAGTCATAGG ATCAATTTT TAAGTACACA TTAGCTGTGA
2160
CTAATGTATA AGAACAACTA TAAAACAAAT AAACAGTGGT T
2201

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 451 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Thr Lys Arg Lys Leu Arg Asn Asn Trp Ile Ile Val Thr Thr Met
1 5 10 15
Ile Thr Phe Val Thr Ile Phe Leu Phe Cys Leu Ile Ile Ile Phe Phe
20 25 30
Leu Lys Asp Thr Leu His Asn Ser Glu Leu Asp Asp Ala Glu Arg Ser
35 40 45

Ser Ser Asp Ile Asn Asn Leu Phe His Ser Lys Pro Val Lys Asp Ile
 50 55 60
 Ser Ala Leu Asp Leu Asn Ala Ser Leu Gly Asn Phe Gln Glu Ile Ile
 65 70 75 80
 Ile Tyr Asp Glu His Asn Asn Lys Leu Phe Glu Thr Ser Asn Asp Asn
 85 90 95
 Thr Val Arg Val Glu Pro Gly Tyr Glu His Arg Tyr Phe Asp Arg Val
 100 105 110
 Ile Lys Lys Arg Tyr Lys Gly Ile Glu Tyr Leu Ile Ile Lys Glu Pro
 115 120 125
 Ile Thr Thr Gln Asp Phe Lys Gly Tyr Ser Leu Leu Ile His Ser Leu
 130 135 140
 Glu Asn Tyr Asp Asn Ile Val Lys Ser Leu Tyr Ile Ile Ala Leu Ala
 145 150 155 160
 Phe Gly Val Ile Ala Thr Ile Ile Thr Ala Thr Ile Ser Tyr Val Phe
 165 170 175
 Ser Thr Gln Ile Thr Lys Pro Leu Val Ser Leu Ser Asn Lys Met Ile
 180 185 190
 Glu Ile Arg Arg Asp Gly Phe Gln Asn Lys Leu Gln Leu Asn Thr Asn
 195 200 205
 Tyr Glu Glu Ile Asp Asn Leu Ala Asn Thr Phe Asn Glu Met Met Ser
 210 215 220
 Gln Ile Glu Glu Ser Phe Asn Gln Gln Arg Gln Phe Val Glu Asp Ala
 225 230 235 240
 Ser His Glu Leu Arg Thr Pro Leu Gln Ile Ile Gln Gly His Leu Asn
 245 250 255
 Leu Ile Gln Arg Trp Gly Lys Lys Asp Pro Ala Val Leu Glu Glu Ser
 260 265 270
 Leu Asn Ile Ser Ile Glu Glu Met Asn Arg Ile Ile Lys Leu Val Glu
 275 280 285
 Glu Leu Leu Glu Leu Thr Lys Gly Asp Val Asn Asp Ile Ser Ser Glu
 290 295 300
 Ala Gln Thr Val His Ile Asn Asp Glu Ile Arg Ser Arg Ile His Ser
 305 310 315 320
 Leu Lys Gln Leu His Pro Asp Tyr Gln Phe Asp Thr Asp Leu Thr Ser
 325 330 335
 Lys Asn Leu Glu Ile Lys Met Lys Pro His Gln Phe Glu Gln Leu Phe
 340 345 350
 Leu Ile Phe Ile Asp Asn Ala Ile Lys Tyr Asp Val Lys Asn Lys Lys
 355 360 365

Ile Lys Val Lys Thr Arg Leu Lys Asn Lys Gln Lys Ile Ile Glu Ile
370 375 380
Thr Asp His Gly Ile Gly Ile Pro Glu Glu Asp Gln Asp Phe Ile Phe
385 390 395 400
Asp Arg Phe Tyr Arg Val Asp Lys Ser Arg Ser Arg Ser Gln Gly Gly
405 410 415
Asn Gly Leu Gly Leu Ser Ile Ala Gln Lys Ile Ile Gln Leu Asn Gly
420 425 430
Gly Ser Ile Lys Ile Lys Ser Glu Ile Asn Lys Gly Thr Thr Phe Lys
435 440 445
Ile Ile Phe
450

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 736 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATTTACGTTT TGTCACTCGTA TCACATACCC AACGCCACGA ACTGTTCAA TCATTTGTC
60
ACGATCGTAT GGTTTTAACT TGTTTCGTAATATCTTATA TAAACATCTA CGACATTTGT
120
TTCTACTTCA CTATTATAAC CCCATACATG ATTTAAAATT TGTTCCCGTT GCATAACATG
180
GTTTTTATTT TCAGCTAGAA GATATAGTAA ATCATACTCT GTTTTGTTA ATTCAATTTC
240
TGCGCCATTT ACCGTCACTT TAAAAGCGTT CTTATCAATT GTAATACCGT TGACATCGAT
300
AATATCCTTT TGTGGCTGAC GACGTAAAAT TGCACGAATT CTTGCTAAAA GTTCTTCAAT
360
ATCAAACGGC TTAACTATAT AATCGTCTGC ACCGTAATCA AGCCCAGCAA CTTTGTCAATA
420
CGTATCACTT TTGCGCTGTAATGAT AGGTGTAGAT TGTTGTTGTC TAATTTGCG
480

ACAAATTCTT AAGCCATTA TTGACGGCAA CATTAAATCT AATATGATTA AATCATAGTA
540
ATGGCTAAGC GCTTTATCTA AACCGTCTTG TCCATCATACT TCTGTGTCCA CATTGTAATT
600
TTCATGTGTG AGTTCCAATT CAAGAAATCT TGCTAAGTTT TGTTCATCTT CTACTATTAA
660
AATTGCGTC ATTTGTACAC CTCATATTAC GACTTTCT AATAAGGTAA TATATTTAA
720
ATTTGTATT TTTCTA
736

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 219 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Thr Gln Ile Leu Ile Val Glu Asp Glu Gln Asn Leu Ala Arg Phe
1 5 10 15
Leu Glu Leu Glu Leu Thr His Glu Asn Tyr Asn Val Asp Thr Glu Tyr
20 25 30
Asp Gly Gln Asp Gly Leu Asp Lys Ala Leu Ser His Tyr Tyr Asp Leu
35 40 45
Ile Ile Leu Asp Leu Met Leu Pro Ser Ile Asn Gly Leu Glu Ile Cys
50 55 60
Arg Lys Ile Arg Gln Gln Gln Ser Thr Pro Ile Ile Ile Ile Thr Ala
65 70 75 80
Lys Ser Asp Thr Tyr Asp Lys Val Ala Gly Leu Asp Tyr Gly Ala Asp
85 90 95
Asp Tyr Ile Val Lys Pro Phe Asp Ile Glu Glu Leu Leu Ala Arg Ile
100 105 110
Arg Ala Ile Leu Arg Arg Gln Pro Gln Lys Asp Ile Ile Asp Val Asn
115 120 125
Gly Ile Thr Ile Asp Lys Asn Ala Phe Lys Val Thr Val Asn Gly Ala
130 135 140

Glu Ile Glu Leu Thr Lys Thr Glu Tyr Asp Leu Leu Tyr Leu Leu Ala
145 150 155 160
Glu Asn Lys Asn His Val Met Gln Arg Glu Gln Ile Leu Asn His Val
165 170 175
Trp Gly Tyr Asn Ser Glu Val Glu Thr Asn Val Val Asp Val Tyr Ile
180 185 190
Arg Tyr Leu Arg Asn Lys Leu Lys Pro Tyr Asp Arg Asp Lys Met Ile
195 200 205
Glu Thr Val Arg Gly Val Gly Tyr Val Ile Arg
210 215

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATGACAAAAC GTAAATTGCG CAATAAC

27

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AAATATGATT TTAAACGTTG TTCC

24